Tracking of Non-rigid Targets in 3D US Images: Results on CLUST 2015

Lucas Royer[1,2,3], Guillaume Dardenne[1], Anthony Le Bras[1,4], Maud Marchal[1,3], Alexandre Krupa[1,2]

[1] Institut de Recherche Technologique b-com, Rennes, France
[2] Inria Rennes - Bretagne Atlantique, France
[3] INSA de Rennes, France
[4] CHU de Rennes, France

Abstract. In ultrasound-guided procedures, such as high-intensity focused ultrasound or radio-frequency ablation, non-rigid clinical targets may undergo displacements due to physiological motions. To cope with that issue, the accurate estimation of the target motion is required in order to adjust the position of medical tools. In this paper, we present a robust approach that allows to track in real-time deformable targets in 3D ultrasound images. Our method combines visual motion estimation with a mechanical model of the target. Our approach is evaluated on the MICCAI CLUST'15 challenge 3D database. We achieved a mean tracking error of 1.78 mm with an average computation time of 350 ms per frame.

1 Introduction

Over the last few years, minimally-invasive procedures, such as high-intensity focused ultrasound (HIFU), or radio-frequency ablation (RFA) have gained significant attention due to the shorter recovery time compared to conventional therapies. However, the quality of these therapies can strongly depend on both the deformations and displacements of the clinical targets since the surgeon needs to continuously adjust the positions of medical tools. Thus, to ensure the target visibility under ultrasound (US) image guidance, several target tracking methods have been developed [8–14]. We recently presented in [7] a robust approach that combines a mechanical model and visual estimation. The good performance of this method has been showed on data obtained from a deformable soft tissue phantom. In this paper, we demonstrate that this method can achieve high accuracy on real-data by testing our algorithm on the database proposed by the MICCAI CLUST'15 challenge. The rest of the paper is organized as follows. In section 2, we present the method that allows to track deformable target in 3D ultrasound images. In section 3, we describe the performance of our approach on real-data. Finally, section 4 concludes the paper.

2 Method

The objective of our approach is to track the motions of a clinical target in 3D ultrasound sequence. To do so, we manually segment the target within the first 3D ultrasound image and we generate a corresponding 3D tetrahedral mesh model (section 2.1). Once the model is defined, we need to estimate the target motions over the consecutive frames. For this purpose, the model displacements are obtained by iteratively summing the internal displacements estimated from a mechanical component (section 2.2), and the external displacements computed from an intensity-based approach (section 2.3).

2.1 Piece-wise Affine Model

In 3D US images, a clinical target can be represented by a continuous set of N_v voxels that is delimited by a visible border. Typical examples are shown in Fig. 2. In order to define the target, we first extract its shape at the initial 3D frame of the US sequence by performing a segmentation. To remove sharp edges and discontinuous shapes, a smoothing step is performed on the 3D segmented surface and a corresponding fitted tetrahedral mesh containing N_c vertices is defined. Then, in order to represent the displacements of the voxels, we propose to use a piece-wise affine warp function. Our piece-wise affine warping is parameterized from both the vertice positions and an affine interpolation that uses barycentric coordinates as proposed in [15]. In this way, we can relate all the voxel positions with all the vertices as follows:

$$\mathbf{p}_{im} = \mathbf{M}.\mathbf{q} \tag{1}$$

where $\mathbf{M}$ is a $(3 \cdot N_v) \times (3 \cdot N_c)$ constant matrix defining the set of barycentric coordinates. $\mathbf{p}_{im}$ is a $(3 \cdot N_v) \times 1$ vector defining all the voxels positions, and $\mathbf{q}$ is a $(3 \cdot N_c) \times 1$ vector containing all the vertices positions. Thanks to Eq. (1), we can update the positions of the target when the vertices of the model are displaced. To compensate the lack of smoothness as well as the poor estimation of vertice positions in US images, we combine a mechanical model to the estimation of displacement.

2.2 Mechanical Component

Our approach combines a mass-spring-damper system to the mesh model previously described. Thus, the vertice displacements are constrained by linking each connected vertice pair with a spring ensuring physically-plausible and coherent displacements of the vertices. Furthermore, the mass-spring-damper system can be specifically characterized by setting a mass value to each vertex, together with elastic and damping coefficients on each spring depending on the soft-tissues homogeneities. In this work, these values have been set empirically but another solution could consist in estimating these parameters from elastography images. However, a further study is required in order to define the impact

on the accuracy of the approach. From this model, we can compute the force $\mathbf{f}_{ij} = [f_{x_{ij}}\ f_{y_{ij}}\ f_{z_{ij}}]^T$ applied on a vertice $\mathbf{q}_i$ from a neighbor vertice $\mathbf{q}_j$. This force can be expressed as follows:

$$\mathbf{f}_{ij} = K_{ij}(d_{ij} - d_{ij}^{init})(\mathbf{q}_i - \mathbf{q}_j) + D_{ij}(\dot{\mathbf{q}}_i - \dot{\mathbf{q}}_j) \circ (\mathbf{q}_i - \mathbf{q}_j) \tag{2}$$

where d_{ij} and d_{ij}^{init} respectively represent the distance between the vertices $\mathbf{q}_i$ and $\mathbf{q}_j$ at their current positions and at their initial positions. The $\circ$ operator expresses the Hadamard product, K_{ij} is a scalar value denoting the stiffness of the spring that links the two vertices while D_{ij} is the damping coefficient value. By combining the previous equation for all the vertices, we can express the total amount of forces $\mathbf{f}_i$ exerted on each vertice $\mathbf{q}_i$ of the mesh model as follows:

$$\mathbf{f}_i = \sum_{n=0}^{N_i} \mathbf{f}_{in} + G_i \dot{\mathbf{q}}_i \tag{3}$$

N_i denotes the number of neighbors vertices connected to the vertice $\mathbf{q}_i$. G_i represents the velocity damping coefficient associated to the vertice $\mathbf{q}_i$. In order to obtain the displacements $\Delta\mathbf{d}$ associated to the mass-spring-damper system, we integrate the forces expressed in Eq. (3) with a semi-implicit Euler integration scheme. Such mechanical constraint can ensure the smoothness warping function of the deformation and limits the noise sensitivity of the intensity-based approach.

2.3 Additive Update Tracking

Let us recall that the main objective of our approach is to estimate the new positions of the target in 3D US sequence. To do so, we use an intensity-based method that consists in evaluating the vertice displacements by minimizing a dissimilarity function E. Therefore, we can express the cost function which minimizes image dissimilarity from the relationship described in Eq. (1) such that:

$$C(\Delta\mathbf{q}) = E(I_t(\mathbf{p}_{im}(t)), I_{t_0}(\mathbf{p}_{im}(t_0))) = E(I_t(\mathbf{M}\mathbf{q}(t)), I_{t_0}(\mathbf{M}\mathbf{q}(t_0))) \tag{4}$$

where I_{t_i} is a vector representing the US intensity of the volume acquired at time index t_i. $\Delta\mathbf{q}$ denotes the vertices displacements. $\mathbf{p}_{im}(t_i)$ and $\mathbf{q}(t_i)$ represent respectively the voxel positions and the vertice positions at time index t_i. In order to determine the dissimilarity function E, we assume that the intensity values of soft tissues are consistent over the time. Consequently, we propose to use the Sum of Squared Differences (SSD) in order to measure the image error. The cost function can now be expressed as:

$$C(\Delta\mathbf{q}) = (I_t(\mathbf{M}(\mathbf{q}(t))) - I_{t_0}(\mathbf{M}(\mathbf{q}(t_0))))^2 \tag{5}$$

The objective is to find iteratively the vertice displacements by minimizing the cost function C. To do so, we perform a Taylor expansion of the previous equation that leads to:

$$C(\Delta\mathbf{q}) \approx (\mathbf{J}\Delta\mathbf{q} + I_t(\mathbf{M}(\mathbf{q}^{k-1}(t))) - I_{t_0}(\mathbf{M}(\mathbf{q}(t_0))))^2 \tag{6}$$

where $\mathbf{q}^{k-1}(t)$ represents the estimation of the parameters at time t at iteration $k-1$ of the optimization algorithm. $\mathbf{J}$ denotes the Jacobian matrix associated to the cost function. This matrix relates the variation of the parameters $\Delta\mathbf{q}$ with the intensity variation of I. It can be computed as follows:

$$\mathbf{J} = \nabla I.\mathbf{M} \tag{7}$$

where ∇I denotes the gradient of the current 3D US frame. In order to obtain the optimal displacements of the vertices, we chose to use a forward-additive steepest gradient strategy as it is computationally efficient since it does not require the calculation of pseudo-inverse of large Jacobian matrix. We therefore obtain:

$$\Delta\mathbf{q} = -\alpha\mathbf{J}^t[I_t(\mathbf{M}(\mathbf{q}^{k-1}(t))) - I_{t_0}(\mathbf{M}(\mathbf{q}(t_0)))] \tag{8}$$

where $\alpha > 0$ denotes the step size of the minimization strategy. $\mathbf{J}^t$ represents the transpose matrix of the Jacobian $\mathbf{J}$. As stated previously, in order to prevent inaccurate results, we propose to combine this motion estimation with the internal displacements of the mass-spring-damper system. This can be performed by iteratively estimate the optimal displacement as follows:

$$\mathbf{q}_k(t) = \mathbf{q}_{k-1}(t) + \Delta\mathbf{q} + \Delta\mathbf{d} \tag{9}$$

where $\Delta\mathbf{d}$ is the internal displacements obtained from the integration of forces expressed in Eq. (3). $\Delta\mathbf{q}$ represents the external displacements from the steepest gradient strategy in Eq. (8). $\mathbf{q}_{k-1}(t)$ denotes the estimation of vertice position at iteration $k-1$ and at time index t. In order to balance the influence of the mass-spring-damper system regarding to the motion estimation between $\Delta\mathbf{q}$ and $\Delta\mathbf{d}$, we can tune the α coefficient that represents the step size of the minimization strategy in Eq.(8).

3 Results

3.1 Description of our Evaluation Environment

Our approach has been tested on real-data and has been implemented with C++/GPU code by using Cuda and VTK libraries. The segmentation step in the first volume is performed with the ITK-SNAP [16] software and can be executed in less than 3 minutes. The mesh is generated thanks to the tetGen [17] software. The resulted computation time of the online tracking is 350 ms allowing thus real-time capabilities. The code was executed on a Windows 7 machine with an Intel core i7-3840qm(2.80GHz).

3.2 Validation Results on Real-data

In order to evaluate our method, we used the database provided by the workshop MICCAI CLUST'15 challenge. The main goal of this challenge is to compare

different state-of-the-art methods for tracking anatomical landmarks in US sequences. For this purpose, a database containing 2D/3D ultrasound sequences of volunteers under free breathing is provided. Furthermore, in order to generate ground truth data, the positions of the target landmarks are identified from expert annotations for each frame. Thus, a comparison can be performed between the ground truth landmark positions and the point position estimated from the tracking task over each frame. The point position at each frame is retrieved by using the equation (1). Is is worth mentioning that both the ultrasound sequences and the annotations are provided by several research institutes. Thus, the approach has been tested by tracking 32 different anatomical features acquired from 16 3D US sequences. For each experiment, we set the elastic and damping parameters such that $K_{ij} = 3.0$ and $D_{ij} = 0.1$ for all the springs, along with $G_i = 2.7$ for all vertices. These values have been set empirically by comparing the tracking accuracy with different set of parameters. This evaluation has been performed on the training database of MICCAI CLUST'15 challenge. The step size of the steepest gradient method has been set to $\alpha = 2 \times 10^{-6}$. In future work, we plan to automatically estimate these parameters by using elastography images. The preliminary results are reported in Fig. 1. From this figure, we can

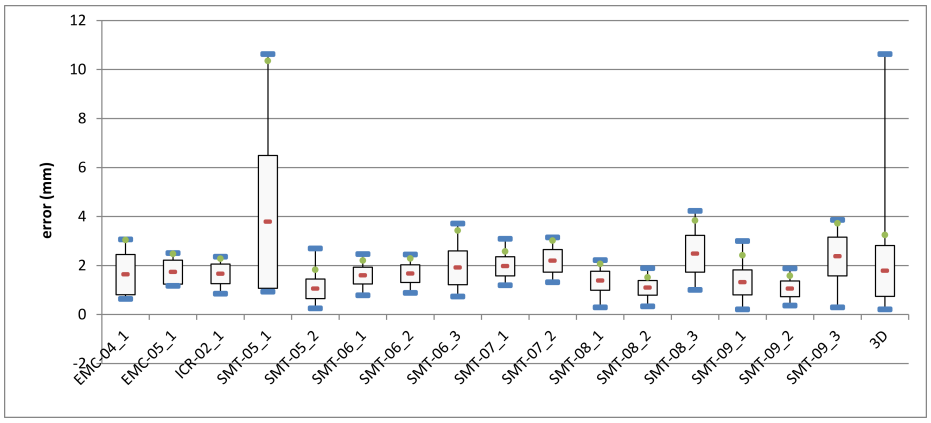

Fig. 1. Tracking error results for each tracking task. The x-axis and y-axis represent respectively each ultrasound sequence and the associated tracking error expressed in millimeters. The name of the sequence (e.g. SMT-05_1) represent both the acronym of the institute and the sequence index. The right box plot "3D" represents the average results for all the sequences (Red) Mean tracking error estimated from euclidean distance. (Black box) Mean error $\pm$ standard deviation. (Whiskers) Minimum and maximum errors. (Green dot) 95^{th} percentile of error.

notice that our approach performs accurate tracking since the mean tracking error is under 2.5 mm for most of the ultrasound sequences. However, we can observe that we obtained some unsatisfactory results (SMT-05_01) when the target goes out of the field of view (FOV). In this case, the error is introduced by

the voxels that are not within the field of view since they have a strong negative influence in the cost function (Eq. 5). This issue can be tackled by using prescan data as suggested in [7]. The performance of our approach is also illustrated in Fig. 2 showing the tracking results at different frames on four landmarks representing hepatic vein bifurcations. In addition, we can notice that our method remains robust with empirical parameters.

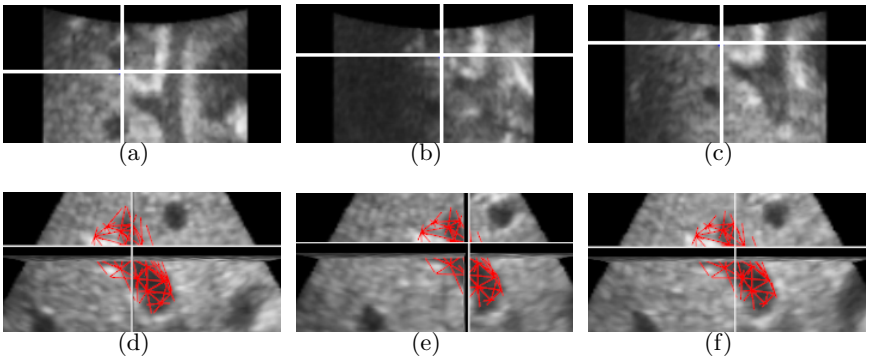

Fig. 2. Example of the tracking task on two sequences. (a-b-c) Tracking of landmark representing portal vein birfucation at frame index 00 (a), 23 (b), 59 (c). (d-e-f) Tracking of landmark representing first degree bifurcation of hepatic bile duct. The red model represents the associated 3D mesh model at frame index 00 (d), 28 (e), 78 (f).

4 Conclusion

In this paper, we presented a method for tracking and automatically compensating the displacements of a deformable target in 3D ultrasound images. The robustness of our tracking method is ensured by combining a mechanical model to the displacement estimation. We evaluated the good performance of our approach through CLUST'15 challenge database. In future work, we plan to automatically estimate the elastic parameters by using elastography images.

References

1. D. Angelova and L. Mihaylova, "Contour segmentation in 2d ultrasound medical images with particle filtering," *Machine Vision and Applications*, vol. 22, no. 3, pp. 551–561, 2010.
2. F. Yeung, S. F. Levinson, D. Fu, and K. J. Parker, "Feature-adaptive motion tracking of ultrasound image sequences using a deformable mesh," *IEEE Trans. on Medical Imaging*, vol. 17, no. 6, pp. 945–956, 1998.
3. A. Basarab, H. Liebgott, F. Morestin, A. Lyshchik, T. Higashi, R. Asato, and P. Delachartre, "A method for vector displacement estimation with ultrasound imaging and its application for thyroid nodular disease," *Medical Image Analysis*, vol. 12, no. 3, pp. 259–274, June 2008.

4. I. Mikic, S. Krucinski, and J. D. Thomas, "Segmentation and tracking in echocardiographic sequences: active contours guided by optical flow estimates," *IEEE Trans. on Medical Imaging*, vol. 17, no. 2, pp. 274–284, 1998.
5. D. Lee and A. Krupa, "Intensity-based visual servoing for non-rigid motion compensation of soft tissue structures due to physiological motion using 4d ultrasound," in *Proc. of IEEE International Conference on Intelligent Robots and Systems*, 2011, pp. 2831–2836.
6. O. Somphone, S. Allaire, B. Mory, and C. Dufour, "Live Feature Tracking in Ultrasound Liver Sequences with Sparse Demons," in *Proc. of MICCAI Workshop on Challenge on Liver Ultrasound Tracking*, 2014, p. 53.
7. L. Royer, M. Marchal, A. Le Bras, G. Dardenne, and A. Krupa, "Real-time Tracking of Deformable Target in 3d Ultrasound Images," in *IEEE Int. Conf. on Robotics and Automation*, 2015.
8. V. De Luca, M. Tschannen, G. Szkely, and C. Tanner, "A learning-based approach for fast and robust vessel tracking in long ultrasound sequences," in *Medical Image Computing and Computer-Assisted InterventionMICCAI 2013*. Springer, 2013, pp. 518–525.
9. V. De Luca, T. Benz, S. Kondo, L. Knig, D. Lbke, S. Rothlbbers, O. Somphone, S. Allaire, M. A. Lediju Bell, D. Y. F. Chung, A. Cifor, C. Grozea, M. Gnther, J. Jenne, T. Kipshagen, M. Kowarschik, N. Navab, J. Rhaak, J. Schwaab, and C. Tanner, "The 2014 liver ultrasound tracking benchmark," *Physics in Medicine and Biology*, vol. 60, no. 14, pp. 5571–5599, July 2015.
10. F. Preiswerk, V. De Luca, P. Arnold, Z. Celicanin, L. Petrusca, C. Tanner, O. Bieri, R. Salomir, and P. C. Cattin, "Model-guided respiratory organ motion prediction of the liver from 2d ultrasound," *Medical Image Analysis*, vol. 18, no. 5, pp. 740–751, July 2014.
11. M. Lediju, B. C. Byram, E. J. Harris, P. M. Evans, J. C. Bamber, and others, "3d Liver tracking using a matrix array: Implications for ultrasonic guidance of IMRT," in *Ultrasonics Symposium (IUS), 2010 IEEE*. IEEE, 2010, pp. 1628–1631.
12. M. Lediju Bell, B. Byram, E. Harris, P. Evans, and J. Bamber, "In vivo liver tracking with a high volume rate 4d ultrasound scanner and a 2d matrix array probe," *Physics in Medicine and Biology*, 2012.
13. S. Vijayan, S. Klein, E. F. Hofstad, F. Lindseth, B. Ystgaard, and T. Lango, "Validation of a non-rigid registration method for motion compensation in 4d ultrasound of the liver," in *Biomedical Imaging (ISBI), 2013 IEEE 10th International Symposium on*. IEEE, 2013, pp. 792–795.
14. J. Banerjee, C. Klink, E. D. Peters, W. J. Niessen, A. Moelker, and T. van Walsum, "4d Liver Ultrasound Registration," in *Biomedical Image Registration*, D. Hutchison, T. Kanade, J. Kittler, J. M. Kleinberg, A. Kobsa, F. Mattern, J. C. Mitchell, M. Naor, O. Nierstrasz, C. Pandu Rangan, B. Steffen, D. Terzopoulos, D. Tygar, G. Weikum, S. Ourselin, and M. Modat, Eds. Cham: Springer International Publishing, 2014, vol. 8545, pp. 194–202.
15. I. Matthews and S. Baker, "Active appearance models revisited," *International Journal of Computer Vision*, vol. 60, no. 2, pp. 135–164, 2004.
16. P. A. Yushkevich, J. Piven, H. C. Hazlett, R. G. Smith, S. Ho, J. C. Gee, and G. Gerig, "User-guided 3d active contour segmentation of anatomical structures: Significantly improved efficiency and reliability," *NeuroImage*, vol. 31, no. 3, pp. 1116–1128, July 2006.
17. H. Si, "TetGen, a Delaunay-Based Quality Tetrahedral Mesh Generator," *ACM Trans. on Mathematical Software*, vol. 41, no. 11, 2015.